AD-A 03 813

TECHNICAL REPORT ARBRL-TR-02343

VARIABLE TRANSFORMATION IN NONLINEAR LEAST SQUARES MODEL FITTING

Aivars Celmiņš

July 1981

US ARMY ARMAMENT RESEARCH AND DEVELOPMENT COMMAND
BALLISTIC RESEARCH LABORATORY
ABERDEEN PROVING GROUND, MARYLAND

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER TECHNICAL REPORT ARBRL-TR-02343	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Variable Transformation in Nonlinear Least Squares Model Fitting		5. TYPE OF REPORT & PERIOD COVERED Final 6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) Aivars Celmiņš		8. CONTRACT OR GRANT NUMBER(s)
9. PERFORMING ORGANIZATION NAME AND ADDRESS U.S. Army Ballistic Research Laboratory ATTN: DRDAR-BLI Aberdeen Proving Ground, MD 21005		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 1L161102AH43
11. CONTROLLING OFFICE NAME AND ADDRESS U.S. Army Armament Research & Development Command U.S. Army Ballistic Research Laboratory ATTN: DRDAR-BL Aberdeen Proving Ground, MD 21005		12. REPORT DATE JULY 1981 13. NUMBER OF PAGES 47
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) Unclassified 15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Nonlinear Least Squares
Model Fitting
Nonlinear Transformations
Iteration Algorithms
Variance Estimates

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)(meg/dll)

The numerical treatment of nonlinear model fitting problems often can be simplified by manipulating the model equations. Algebraic manipulations, including nonlinear transformations of model parameters, do not change the numerical result of the adjustment. Therefore, such manipulations can be a powerful method to improve the performance of solution algorithms. Nonlinear transformations of the observations, on the other hand, do change the numerical results unless the normal equations are transformed accordingly. (Cont'd)

20. Abstract (Cont'd):

The latter transformation has been neglected by previous authors and this article provides a complete set of formulas that are needed to implement transformations of observations. The transformations are, however, in general less useful than parameter transformations but may have applications in particular situations.

TABLE OF CONTENTS

I. INTRODUCTION

A mathematical model fitting problem arises when one compares real observations with theoretical predictions. The observations always contain observational inaccuracies and, likewise, the theory of the prediction can be inadequate. If discrepancies between observations and predictions are unacceptably large for a particular situation, then one is faced with the task to adjust in a rational manner either the observations, or the theory, or both so that an acceptable mathematical description of the event can be established. The problem can be subdivided conveniently into three subtasks, each of which requires a different approach and background information.

First, one has to chose a model. Normally, this requires supporting information from engineering, physics, geometry, etc., which may suggest or postulate a reasonable mathematical description of the observable event. We shall assume in this report that the model is formulated as a system of equations containing observations and, possibly, also some undetermined model parameters.

Once the model is selected, one can compare predicted values of observable quantities with corresponding observations. The comparison provides the basis for a rational adjustment of the observations and/or of the model. This subtask of the problem is a purely mathematical part of model fitting and it belongs to the category of ill-posed problems. Its mathematical/numerical treatment is independent of the other two subtasks, i.e., of applications. We shall be concerned with this part of the problem in the present article.

After the adjustments have been carried out, one has to validate the mathematical model, unless it has been prescribed, e.g., by the geometry of the event. The validation involves typically, but not necessarily, a statistical analysis of the discrepancies between observations and predictions. The result of the validation process may be a new formulation of model equations and subsequent fitting, i.e., a repetition of the whole task until some validation criterion is satisified. We shall not discuss this part of the problem, noticing only that the results of the second subtask provide the data basis necessary for a validation.

If the model equations are not linear then the model fitting problem generally leads to systems of complicated simultaneous equations and corresponding numerical difficulties may arise. Often the numerical treatment can be simplified by a reformulation of the model equations, particularly by introduction of new variables through variable transformations. Such manipulations have been suggested in textbooks[1-7] and are routinely used in applications. Examples of recently published applications where variable transformations have been used are References 8, 9, and 10.

A closer investigation of variable transformations in model fitting problems suggests that the formulations should be used more cautiously than some of the texts suggest. Therefore, we shall, in this report, present an investigation of some consequences of the transformations and draw conclusions about their usefulness for the simplification of the numerical treatment of model fitting problems.

In Section II we shall formulate the mathematical model fitting problem in general terms and discuss the effects that can be anticipated from manipulations of model equations. In Section III we shall specialize the considerations to nonlinear least squares problems and produce explicit formulas that are needed in such problems. Some examples will be presented in Section IV, and Section V will summarize the conclusions that can be drawn from the theoretical discussions and from examples.

---

[1] *W. Edwards Deming, Statistical Adjustment of Data, John Wiley and Sons, London, 1944.*

[2] *John E. Freund, Mathematical Statistics, Prentice Hall, Englewood Cliffs, NJ, 1962.*

[3] *Karel Rektorys, ed., Survey of Applicable Mathematics, The M.I.T. Press, Cambridge, MA, 1969.*

[4] *Yonathan Bard, Nonlinear Parameter Estimation, Academic Press, New York, NY, 1974.*

[5] *Yo Lun Chou, Statistical Analysis, Holt, Rinehart, and Winston, New York, NY, 1975.*

[6] *A.S.C. Ehrenberg, Data Reduction, John Wiley and Sons, New York, NY, 1975.*

[7] *Thomas H. Wonnacott and Ronald F. Wonnacott, Introductory Statistics, 3rd ed., John Wiley and Sons, New York, NY, 1977.*

[8] *R.M. Passi, "Use of Nonlinear Least Squares in Meteorological Applications," Journal of Applied Meteorology, Vol. 16, pp. 827-832, 1977; and Vol. 17, pp. 1579-1580, 1978.*

[9] *Heike von Benda, "Zur Gitterkonstantenbestimung mit Ausgleichsmetoden," Zeitschrift für Kristallographie, Vol. 149, pp. 205-209, 1979.*

[10] *Ph. Wahl, "Analysis of Fluorescence Anistrophy Decays by a Least Square Method," Biophysical Chemistry, Vol. 10, pp. 91-104, 1979.*

II. GENERAL ASPECTS OF MATHEMATICAL MODEL FITTING

Let the model equations be

$$A(X)\theta = 0 \ , \tag{1}$$

where $X \varepsilon R^n$ is the vector of all observations, $\theta \varepsilon R^p$ is a model parameter vector, and $A(X)$ is an operator that operates on θ and has a range R^r. We assume that the following relations hold between the dimensions n, r, and p:

$$n \geq r > p \geq 0 \ . \tag{2}$$

By permitting the dimension p to be zero, we include in our considerations also cases in which the model equations do not contain free parameters. Then Eq. (1) reduces to $A(X) = 0$.

Typical for applications are cases in which the r Eqs. (1) for θ are independent and, because of Eq. (2), do not have a solution. Then one replaces the model equations by another system

$$\tilde{A}(X)\theta = 0 \ , \tag{3}$$

chosing the operator $\tilde{A}(X)$ such that it approximates $A(X)$ and has a solution. The determination of $\tilde{A}(X)$ can be considered as the central part of the model fitting problem.

In order to have a measure for the approximation, we introduce a metric for the operators. Let $\rho[\tilde{A}(X), A(X)]$ be a metric. Then one can formulate the mathematical model fitting as the following constrained minimization problem:

$$\tilde{A}(X)\theta = 0 \ , \quad W\left\{\rho\left[A(X), A(X)\right]\right\} = \min. \ , \tag{4}$$

where $W\{\rho\}$ is a generally convex object function. The choice of the metric ρ and of the object function $W\{\rho\}$ determines the type of the model fitting, e.g., least squares, maximum norm, etc.

We shall now discuss the selection of an approximate operator $\tilde{A}(X)$. First, we notice that the model operators $A(X)$ and $\tilde{A}(X)$ are generally needed and defined only within a finite neighborhood of the observations X. Therefore, assumptions about properties of the operators need to be made for that neighborhood only. Let the neighborhood consist of all points $Z = X + C$, whereby C is restricted component-wise by

$$\gamma_i \leq C_i \leq \Gamma_i, \; i = 1, 2, \ldots, n . \tag{5}$$

The intervals (γ_i, Γ_i) normally contain zero, but exceptions are possible and do occur in applications. Second, we assume that within the neighborhood (5) $A(Z)$ is a continuous function of Z. Then a reasonable choice of $\tilde{A}(X)$ is

$$\tilde{A}(X) = A(X + C) . \tag{6}$$

The choice achieves a natural parametrization of the approximation. The approximation parameter is the vector $C \varepsilon R^n$ and the operator $\tilde{A}(X)$ depends continuously on the parameter within the restrictions (5).

The parametrized model fitting problem can be formulated as follows:

$$\left.\begin{array}{l} A(X + C)\theta = 0 , \\ \\ W\{\rho[A(X + C), A(X)]\} = \min. \end{array}\right\} \tag{7}$$

The quantities to be determined by Eq. (7) are the approximation parameter C and the model parameter θ. We assume that the solution vector C is within the limits specified by Eq. (5).

We will need in the sequel some differentiability properties for the model operator. As far as X is concerned, we assume the properties to hold within the neighborhood (5). With respect to θ we assume that a similar neighborhood exists in the vicinity of the solution of Eq. (7) in which $A(X)\theta$ is a continuous function of θ. The differentiability assumptions are that $A(X + C)\theta$ is twice differentiable with respect to all its $n + p$ arguments within the cartesian product space of the neighborhoods of X and θ. We also assume that within that space

$$\operatorname{rank} \frac{\partial A}{\partial X} = r , \tag{8}$$

and define

$$\rho\left[A(Z),\ A(X)\right] = ||Z - X|| \ . \tag{9}$$

ρ is a metric within the neighborhood in which Eq. (8) holds. We also assume that the model equations do not contain redundant parameters. The assumption may be expressed as the requirement

$$\operatorname{rank} \frac{\partial A(X)\theta}{\partial \theta} = p \ . \tag{10}$$

With the specialization Eq. (9), the model fitting problem becomes

$$\left.\begin{aligned} &A(X + C)\theta = 0 \ , \\ &W\{\rho\left[A(X + C),\ A(X)\right]\} = W\{\,||C||\,\} \ = \min. \end{aligned}\right\} \tag{11}$$

Eq. (11) is an abstract formulation of common model fitting problems. The difference C between the observations X and the "corrected observations" X + C is called the residual vector. In the formulation (11), we require that a norm of the residual vector be minimized, subject to model equations which have to be satisfied at X + C. The model parameter vector θ is not essential in this formulation. The number of model parameters may be zero and it is normally orders of magnitudes smaller than the number of approximation parameters; i.e., residuals. The determination of θ can be, of course, in some applications more important than the determination of C, but this is not always the case.

A least squares model fitting problem is a special case of Eq. (11), characterized by a particular choice of the norm in the definition Eq. (5), and of the object function $W\{\rho\}$. The least squares metric is

$$\rho\left[A(Z),\ A(X)\right] = ||Z{-}X|| = \left[(Z{-}X)^T R^{-1}(Z{-}X)\right]^{1/2} \ , \tag{12}$$

where R is an estimate of the variance-covariance matrix of the observations. The least squares object function is

$$W\{\rho\} = \rho^2 \ . \tag{13}$$

Therefore, the least squares model fitting problem is defined by

$$\left.\begin{aligned} &A(X + c)t = 0 \ , \\ &W = ||c||^2 = c^T R^{-1} c = \min. \end{aligned}\right\} \qquad (14)$$

In Eq. (14) we have used c and t instead of C and θ, respectively, thus indicating the least squares values of both parameter vectors.

The use of R^{-1} as a norm matrix in the definition (11) makes the norm $||C||$ and W dimensionless, which is very convenient when fitting results are compared. If the variance-covariance matrix R is known exactly, and the observational errors are normally distributed, then the solution of Eq. (14) is a maximum likelihood solution of the approximation problem[11]. The same maximum likelihood solution is obtained if R approximates the variance-covariance matrix up to an unknown factor. In applications one has to be content with an estimate of R. Then, often, the off-diagonal elements are assumed to be zero as a matter-of-course. Because the results of the model fitting depend on R, such assumptions should not be made without having reasons that zero is a better approximation than a non-zero value. The theoretical treatment is not complicated by the assumption that R is not diagonal, nor are the numerical complications unsurmountable. Realistic estimates of R are, however, important for the interpretation of the results, and for the validation of the fitting.

We solve the optimization problem (11) or (14) using Lagrange multiplier technique, and call the multipliers correlates, as usual in adjustment problems. Let $K \varepsilon R^r$ be a correlate vector and let the modified object function be

$$\tilde{W} = \frac{1}{2} W\{||C||\} - K^T A(X + C)\theta \ . \qquad (15)$$

Necessary conditions for the solution of the optimization problem are obtained by setting equal to zero the partial derivatives of $\tilde{W}$ with respect to the unknown C, θ, and K. This yields the following set of normal equations.

---

[11] *H.J. Britt and R.H. Luecke, "The Estimation of Parameters in Nonlinear, Implicit Models,"* <u>*Technometrics*</u>*, Vol. 15, pp. 233-247, 1973.*

$$\frac{1}{2}\frac{\partial}{\partial C} W\{||C||\} - \frac{\partial}{\partial C}[K^T A(X + C)\theta] = 0 , \tag{16a}$$

$$\frac{\partial}{\partial \theta}[K^T A(X + C)\theta] = 0 , \tag{16b}$$

$$A(X + C)\theta = 0 . \tag{16c}$$

The solution of the model fitting problem (11) is among the solutions of Eqs. (16). On the other hand, one cannot guarantee that a particular solution of the normal equations corresponds to the absolute minimum solution of Eq. (11), nor is the uniqueness of the solution given. An investigation of these complications is not the subject of this paper. Mostly, such problems can be, and are taken care of by ad hoc measures based on background information from the application. Therefore, we simplify our present theoretical discussion by assuming in this section that a numerical solution of Eqs. (16) can be obtained, and that it has been verified as the absolute minimum solution of Eq. (11).

In least squares problems, the first term $\partial W/\partial C$ in Eq. (16a) is linear with respect to C. Nonlinear expressions which could be possibly simplified by algebraic manipulations may occur in the second term in Eq. (16a), and in Eqs. (16b) and (16c). The structure of these terms strongly depend on the form in which the model Eqs. (16c) are cast, and it is obvious that simplifications can be achieved by proper formulations. Particularly, one does not have to insist that each model equation be solved for a "dependent" observation. Such a form is assumed in most textbooks on data reduction and postulated in computer programs for data reduction problems. Quite often an implicit formulation of the Eqs. (16c) can be simpler, producing also simpler expressions for the derivatives in Eqs. (16a) and (16b). The solution of the problem (11) is, of course, independent of the particular form in which the model equations are cast. This remark is trivial in the present context, and it is a consequence of the formulation of the model fitting problem by Eq. (11). Reference 12 reports about numerous unsuccessful attempts to achieve a similar invariance statement when the problem was formulated differently.

The aforementioned manipulations of the model equations can also include nonlinear transformations of the parameter θ. Such transformations do not affect the definition of the metric ρ, because the metric of the operator is independent of the operand. Therefore, the transformations do not affect the first term in Eq. (16a) either, and are a

---

[12] *P.A.D. DeMaine, "Automatic Curve Fitting, I. Test Methods," Computers and Chemistry, Vol. 2, pp. 1-6, 1978.*

powerful tool for the simplification of the rest of the equations. An example in which nonlinear parameter transformations are used to linearize the model equations is reported in Reference 9. In Section IV we shall give other examples.

The formal procedure of replacing parameters is as follows: Suppose that one wants to replace the parameter θ by σ whereby both parameters are related by a nonsingular function

$$\theta = g(\sigma). \tag{17}$$

(Regularity of the transformation need to be assumed only within a neighborhood of the solution.) Let the model equations be in terms of σ

$$\bar{A}(X)\sigma = 0 . \tag{18}$$

The operator $\bar{A}$ can be obtained from A always by the definition

$$\bar{A}(X)\sigma = A(X)g(\sigma) \quad , \tag{19}$$

however, often one can find other equivalent formulations that are simpler. The metric $\bar{\rho}$ associated with $\bar{A}$ is defined as in Eq. (9)

$$\bar{\rho}\left[\bar{A}(Z), \bar{A}(X)\right] = ||Z - X|| \quad . \tag{20}$$

With this definition and the same object function $W\{\rho\}$ as before one obtains the normal equations

$$\frac{1}{2}\frac{\partial}{\partial\bar{C}} W\left\{||\bar{C}||\right\} - \frac{\partial}{\partial\bar{C}}\left[\bar{K}^T\bar{A}(X + \bar{C})\sigma\right] = 0 \ , \tag{21a}$$

$$\frac{\partial}{\partial\sigma}\left[\bar{K}^T\bar{A}(X + \bar{C})\sigma\right] = 0 \ , \tag{21b}$$

$$\bar{A}(X + \bar{C})\sigma = 0 \ . \tag{21c}$$

The solution vectors of Eqs. (16) and Eqs. (21) are related by

$$C = \overline{C} \, , \quad \theta = g(\sigma) \tag{22}$$

The vectors K and $\overline{K}$ can be computed from these values using formulas given in the next section.

The relation (22) is again a simple consequence of the formulation (11) of the model fitting problem. Benda[9] proves the correspondence (22) for a particular transformation and application and indicates that previous developers of software for such problems were not aware of the relation.

If the solution of the model fitting task has been found from Eq. (21) in terms of σ, but the parameter vector θ is of interest, then one needs also a formula for the accuracy of θ. Let us assume that the solution algorithm for Eq. (21) has provided information about the accuracy of σ in form of an estimate V_σ of the variance-covariance matrix of the components of σ. (In Section III, we shall give formulas for V_σ in least squares problems.) Then an estimate of the variance-covariance matrix V_θ of the components of θ can be obtained by applying the linearized law of variance propagation to the relation (22). The result is

$$V_\theta = \frac{\partial g}{\partial \sigma} V_\sigma \left(\frac{\partial g}{\partial \sigma}\right)^T \, . \tag{23}$$

More complicated are consequences of such manipulations of the model equations that involve transformations of the observations. This is so because the transformations now affect the definition of the norm ρ. Next, we shall consider such transformations.

Let a transformation of observations be

$$Y = v(X) \tag{24}$$

with the inverse

$$X = u(Y) \, .$$

We assume that the transformation is regular within the neighborhood (5), including the solution X + C, and that the function u(Y) is there twice differentiable. The model Eqs. (1) are replaced by equivalent (usually simpler) equations

$$\hat{A}(Y)\theta = 0 \ . \tag{25}$$

The operator $\hat{A}(Y)$ can be obtained, e.g., by the definition

$$\hat{A}(Y)\theta = A(u(Y))\theta \ , \tag{26}$$

but, as in the case of parameter transformations, usually other equivalent formulations can be found that are simpler.

When we formulate the model fitting problem in terms of Y, we have to keep in mind that the goal is to minimize the distance C between the actual observations X and their corrected values X + C. In least squares problems, only such a minimization yields under conditions a maximum likelihood solution. Then the minimization problem (11) is

$$\begin{gathered} Y = v(X) \ , \\ \hat{A}(Y + B)\theta = 0 \ , \\ W\{||u(Y + B) - X||\} = \min. \end{gathered} \tag{27}$$

The normal equations for the problem (27) are

$$\frac{1}{2}\frac{\partial}{\partial B} W\{||u(Y + B) - X||\} - \frac{\partial}{\partial B}\left[\hat{K}^T\hat{A}(Y + B)\theta\right] = 0 \ , \tag{28a}$$

$$\frac{\partial}{\partial \theta}\left[\hat{K}^T\hat{A}(Y + B)\theta\right] = 0 \ , \tag{28b}$$

$$\hat{A}(Y + B)\theta = 0 \ . \tag{28c}$$

The first term in Eq. (28a) is not linear with respect to the unknown B unless the transformation (24) is linear. Therefore, a nonlinear transformation that produces an operator $\hat{A}(Y)$ which is simpler than the original operator $A(X)$, introduces nonlinear terms in Eq. (28a). The new nonlinearities may offset the advantages gained by a simplification of the other terms in the equations.

We shall pursue this point further in the next section and show in detail how the normal equations and algorithms are affected by transformations of observations specifically in least squares problems.

III. LEAST SQUARES MODEL FITTING

We consider in this section the effects of variable transformations on least squares model fitting problems. We shall first derive the basic equations for nonlinear least squares problems in terms of the original observations, and then show how the equations are affected by a transformation of the observations. We simplify our notation by defining a vector function $F(X,\theta)$ by

$$F(X,\theta) = A(X)\theta \quad . \tag{29}$$

Then the model Eq. (1) is

$$F(X,\theta) = 0 \ , \tag{30}$$

and the least squares model fitting problem (14) is

$$\left.\begin{aligned} &F(X + c,t) = 0 \ , \\ &W = ||c||^2 = c^T R^{-1} c = \min. \end{aligned}\right\} \tag{31}$$

In the sequel we will use subscripts to denote derivatives. Also, because derivatives of $F(X + c,t)$, with respect to c, are identical to derivatives with respect to X we shall use the subscript X for both. Thus, e.g.,

$$F_X(X+c,t) = \frac{\partial}{\partial X} F(X+c,t) = \frac{\partial}{\partial c} F(X+c,t)$$

and

$$[K^T F(X + c,t)]_{Xt} = \frac{\partial^2}{\partial X \partial t}[K^T F(X + c,t)] = \frac{\partial^2}{\partial c \partial t}[K^T F(X + c,t)]$$

are matrices with the dimensions rxn and nxp, respectively.

Using this notation, the normal equations corresponding to the problem (31) are

$$R^{-1}c - [k^T F_X(X + c,t)]^T = 0 , \tag{32a}$$

$$k^T F_t(X + c,t) = 0 , \tag{32b}$$

$$F(X + c,t) = 0 . \tag{32c}$$

The normal equations are in general nonlinear with respect to c and t. Therefore, their numerical solution will require some kind of iteration. We obtain second order iteration equations for Eqs. (32) by expanding the normal equations at an approximation to the solution and keeping the linear terms of the expansion. Let the approximation to the solution be C, K, and T, and that of the corresponding corrections be ε, κ, and τ. Then the expansion yields the following Newton equations for the corrections:

$$[I-R(K^T F)_{XX}]\varepsilon - RF_X^T \cdot (K + \kappa) - R(K^T F)_{Xt}\tau = - C , \tag{33a}$$

$$(K^T F)_{tX}\varepsilon + F_t^T \cdot (K + \kappa) + (K^T F)_{tt}\tau = 0 , \tag{33b}$$

$$F_X\varepsilon \qquad\qquad + F_t\tau = - F . \tag{33c}$$

The arguments of F and its derivatives in Eqs. (33) are X + C and T.

Newton-Raphson iteration equations can be established by suitable manipulations of Eqs. (33)[8,13,14,15]. A set of such iteration equations are given in the Appendix. Most authors simplify Eqs. (33) by neglecting all terms that contain second order derivatives[1,11,16,17]. This yields so-called Gauss-Newton procedures that have theoretically only linear convergence and that also may have other peculiarities[13].

The final step in a model fitting problem is to obtain variance estimates of the solution in terms of the estimated variances of the observations. We shall restrict ourselves in this article to the estimation of the accuracies of the least squares value t of the parameter vector, and show how the estimation formulas change due to transformations of observables. We shall use the linearized variance propagation formula for the estimates. Estimates of the accuracies of the corrected observations $x = X + c$ can be obtained by analogous processes.

The formulas can be derived from the linear terms of an expansion of the normal Eqs. (33) at the solution[13]. Let dx, dk, and dt be the differentials of the solution vectors $x = X + c$, k, and t, respectively. Then the expansion yields

$$[(I - Rk^TF)_{XX}]dx - RF_X^T dk - R(k^TF)_{Xt}dt = dX \ , \tag{34a}$$

$$(k^TF)_{tX}dx + F_t^T dk + (k^TF)_{tt}dt = 0 \ , \tag{34b}$$

$$F_X dx \qquad\qquad + F_t dt = 0 \ . \tag{34c}$$

The arguments of F and its derivatives in Eqs. (34) are x and t.

---

[13] *Allen J. Pope, "Two Approaches to Nonlinear Least Squares Adjustments," The Canadian Surveyor, Vol. 28, pp. 663-669, 1974.*

[14] *Robert E. Barieau and B.J. Dalton, "Nonlinear Regression and the Principle of Least Squares," Bureau of Mines Report of Investigations 6900, 1967.*

[15] *Aivars K. Celmiņš, "A Manual for General Least Squares Model Fitting," USA ARRADCOM/Ballistic Research Laboratory Report ARBRL-TR-02167, 1979.* (AD #B040229L)

[16] *E. Stark and E. Mikhail, "Least Squares and Non-Linear Functions," Photogrammetric Engineering, pp. 405-412, 1973.*

[17] *William H. Sachs, "Implicit Multifunctional Nonlinear Regression Analysis," Technometrics, Vol. 18, pp. 161-173, 1976.*

By manipulations of Eqs. (34) that can be done in various ways[13,18] one obtains linear relations between dt and dX, and between dx and dX, respectively. Let the former relation be

$$N\, dt = S\, dX \ . \tag{35}$$

(Explicit formulas for N and S are given in the Appendix.) Then the estimated variance-covariance matrix V_t of the parameter vector t is

$$V_t = N^{-1} S\, R\, S^T (N^{-1})^T \ . \tag{36}$$

It is obvious from the derivation of Eq. (36) that V_t, which itself is only a linearized approximation, depends on second order derivatives of F. (The formulas in the Appendix show explicitly this dependency.) Neglect of the second order derivative terms renders a formula that is theoretically less than first order accurate. Therefore, such a neglect has to be justified in each application by providing estimates of the magnitudes of the neglected terms. Of the cited references, complete first order formulas are used only in References 13, 14, 15, and 18.

Next, we introduce variable transformations into the least squares model fitting problem. We can restrict ourselves to transformations of observations because, as shown in Section II, transformations of model parameters have the same effects as simple algebraic manipulations of the model equations.

Let, as in Section II, the transformation be given by

$$Y = v(X) \tag{37}$$

with the inverse

$$X = u(Y) \ .$$

In terms of Y, the least squares model fitting problem is defined by

---

[18] *Aivars Celmiņš, "Least Squares Adjustment with Finite Residuals for Non-Linear Constraints and Partially Correlated Data," Ballistic Research Laboratory Report BRL-R-1658, 1973.* (AD #766283)

$$Y = v(X) \;, \tag{38a}$$

$$H(Y + b, t) = 0 \;, \tag{38b}$$

$$W = ||u(Y + b) - X||^2 = [u(Y + b) - X]^T R^{-1} [u(Y + b) - X] = \text{min.} \tag{38c}$$

Eq. (38b) is a model equation, equivalent to Eq. (30) and expressed in terms of Y.

The normal equations for the problem (38) are

$$[u_Y(Y + b)]^T R^{-1} [u(Y + b) - X] - [k^T H_Y(Y + b, t)]^T = 0 \tag{39a}$$

$$k^T H_t(Y + b, t) = 0 \;, \tag{39b}$$

$$H(Y + b, t) = 0 \;.. \tag{39c}$$

Corresponding Newton equations for corrections β, κ, and τ of approximate solutions B, K, and T, respectively, are

$$[I - Q\Xi]\beta - QH_Y^T \cdot (K + \kappa) - Q(K^T H)_{Yt}\tau = -\Delta \;, \tag{40a}$$

$$(K^T H)_{tY}\beta + H_t^T \cdot (K + \kappa) + (K^T H)_{tt}\tau = 0 \;, \tag{40b}$$

$$H_Y\beta \qquad\qquad + H_t\tau = -H \;, \tag{40c}$$

where

$$Q = v_X R v_X{}^T = (u_Y)^{-1} R \, (u_Y^T)^{-1} \;, \tag{41}$$

$$\Delta = v_X \cdot [u(Y + B) - X] = v_X \cdot C = (u_Y)^{-1} \cdot C \;, \tag{42}$$

$$\Xi = (K^T H)_{YY} - (u^T R^{-1} C)_{YY} \; . \qquad (43)$$

The arguments of the functions H and u in Eqs. (40) through (43) are $Y + B$ and T, and the last term in Eq. (43) is differentiated assuming $C = u(Y + B) - X$ to be constant. The term is a symmetric nxn matrix containing second order derivatives of the transformation function u(Y).

A comparison of Eqs. (40) with Eqs. (33) shows that the important changes in the Newton equations due to the transformation (37) are in Eqs. (40a). The rest of Eqs. (40) is formally identical to the corresponding terms in Eqs. (33), if $F(X,\theta)$ is replaced by $H(Y,\theta)$. In Eqs. (40a) we see three other replacements: the estimated variance-covariance matrix R is replaced by Q, the right-hand side $-C$ is replaced by $-\Delta$, and the term $(K^T F)_{XX}$ is replaced by Ξ.

The replacement of R by Q corresponds to an application of the linearized variance propagation formula to the transformation (37). The replacement of the right-hand sides is a linearized transformation of the residuals C into the Y-space. If the transformation (37) is linear, then only these two replacements occur. If, however, the transformation is nonlinear, then the last term in Eq. (43) does not vanish and, because it contains second order derivatives of u(Y), it can be quite complicated. This complication can offset algorithmic advantages gained by a simplification of other terms in the Newton equations.

Iteration algorithms and formulas for the variances of the solution again can be obtained by manipulations of the Newton equations. Explicit formulas are given in the Appendix. We notice that second order Newton-Raphson algorithms necessarily contain second order derivatives of the model function H as well as of the transformation function u(Y). The coding of the second order derivatives can, of course, be avoided if first order Gauss-Newton algorithms are used. However, variance estimates of the solution can be calculated to a first order accuracy only if all the second order derivatives are available.

The author has carried out numerical experiments to determine whether a solution of Eqs. (39) instead of Eqs. (32) has algorithmic advantages. The experiments were done with the utility programs described in Reference 15. The programs permit one to carry out the calculations either in terms of X, or in terms of Y, and to use either Newton-Raphson, or Gauss-Newton algorithms. The experiments were inconclusive. In some examples the algorithms converged better when the problem was formulated in X, in other examples a formulation in $Y = v(X)$ produced better algorithms. However, the differences in performance were never significant. This result is in strong contrast to similar experiments involving transformations of parameters. In those experiments, a suitable parameter transformation often had a dramatic effect on the performance of the solution algorithm. Some examples are given in the next section.

Another possible benefit from nonlinear transformations of observations could be a simpler problem formulation. The complexity of the normal equations is thereby of secondary importance, if one uses an available general utility program for their solution. However, the model equations must be made available to the utility program, which means that the equations must be programmed. Then one has the choice to program either the function $F(X,\theta)$ with its first and second order derivatives, or the two functions $H(Y,\theta)$ and $u(Y)$ with their derivatives. If the transformation is nonlinear, then normally the programming of H and u will not be simpler than the programming of F. An exception may be the situation where the same transformation $u(Y)$ (e.g., polar-cartesian) is used for several problems with different model functions $H(Y,\theta)$, so that $u(Y)$ has to be programmed only once.

We may conclude that in general a transformation of observations offers little or no advantages over a formulation of the model equations in terms of the original observations. There are, however, other useful applications of such transformations. First, a graphical display of the results can be clearer in terms of Y than in terms of X. Second, and more importantly, the transformations can be a convenient method to derive a "falsified" problem that can be solved easily and that provides initial approximations to the unknown least squares solution vectors. One can falsify the problem; e.g., by using a nonlinear transformation but linearizing its effects on the problem formulation. A simple and effective falsification is to replace the problem (38) by

$$Y = v(X) \ , \tag{44a}$$

$$H(Y + b,t) = 0 \ , \tag{44b}$$

$$W* = b^T \left[u_Y^T(Y)R^{-1}u_Y(Y)\right]b = \min. \tag{44c}$$

The formulation is identical to the correct formulation (38) only if the transformation is linear, but the normal equations for the false problem (44) are simple:

$$\tilde{Q}^{-1}b - \left[k^T H_Y(Y + b,t)\right]^T = 0 \ , \tag{45a}$$

$$k^T H_t(Y + b,t) = 0 \ , \tag{45b}$$

$$H(Y + b,t) = 0 \ , \tag{45c}$$

where

$$\tilde{Q} = [u_Y(Y)]^{-1} R \, [u_Y^T \, (Y)]^{-1} \, . \tag{46}$$

This system can be much simpler and easier to solve than Eqs. (32) or the equivalent Eqs. (39). Its solution is, however, not the least squares solution but an approximate solution of unknown quality.

Initial approximations to the solution also can be obtained by other falsifications in addition to the one described, or instead of it. Such falsifications are, e.g., assumptions that certain observations are error free, that some correlations are zero, that some model parameters have prescribed values, etc.

IV. EXAMPLES

The first example is a case involving transformation between polar and cartesian coordinates. We shall compare results that are obtained using the approach of the previous section with results that are obtained by following suggestions by other authors. In data processing literature, one finds different suggestions. The simplest one is to treat the problem after transformation as if the transformed quantities were observed. It is clear from the discussions in Section II that such an approach does not produce the least squares solution, i.e., it does not minimize $W\{||C||\}$, even if the transformation is linear. The most sophisticated suggestion[1,8,10] is to apply the transformation (46) to R, i.e., to solve the system (45). As we have seen in the previous section, this approach yields the least squares solution only if the transformation $Y = v(X)$ is linear. The following example illustrates the practical consequences of such a problem falsification.

Let the observations be distances r_i and azimuth angles ϕ_i, and let the model equations represent a straight line in cartesian coordinates. Then the model equations are in terms of the original observations

$$F(r,\phi;a,b) = \begin{cases} r_1 \sin\phi_1 - a - br_1\cos\phi_1 = 0 \\ r_2 \sin\phi_2 - a - br_2\cos\phi_2 = 0 \\ \dots\dots\dots\dots\dots \\ r_n \sin\phi_n - a - br_n\cos\phi_n = 0 \end{cases} \tag{47}$$

The transformation of the observations into cartesian coordinates is

$$\left.\begin{aligned} x_i &= r_i \cos\phi_i \ , \\ y_i &= r_i \sin\phi_i \ , \ i = 1, 2, \ldots, n \ , \end{aligned}\right\} \tag{48}$$

and the model equations are in terms of the transformed observations

$$H(X,Y;a,b) = \begin{cases} y_1 - a - bx_1 = 0 \\ y_2 - a - bx_2 = 0 \\ \ldots\ldots\ldots\ldots\ldots \\ y_n - a - bx_n = 0 \end{cases} \tag{49}$$

The Jacobian matrix of the transformation is

$$J = \begin{pmatrix} J_1 & & & 0 \\ & J_2 & & \\ & & \ddots & \\ 0 & & & J_n \end{pmatrix} \ , \tag{50}$$

where

$$J_i = \frac{\partial(x_i,y_i)}{\partial(r_i,\phi_i)} = \begin{pmatrix} \cos\phi_i & -r_i\sin\phi_i \\ \sin\phi_i & r_i\cos\phi_i \end{pmatrix} \tag{51}$$

We assume for simplicity that all observations are independent with estimated standard errors e_{ri} and $e_{\phi i}$, respectively. Then the estimated variance-covariance matrix R is the diagonal matrix

$$R = \begin{pmatrix} e_{r1}^2 & & & & & & & \\ & e_{\phi 1}^2 & & & & 0 & & \\ & & e_{r2}^2 & & & & & \\ & & & e_{\phi 2}^2 & & & & \\ & & & & \ddots & & & \\ & 0 & & & & & e_{rn}^2 & \\ & & & & & & & e_{\phi n}^2 \end{pmatrix} \tag{52}$$

The transformed variance-covariance matrix $\tilde{Q}$ is according to Eq. (46) the block diagonal matrix

$$\tilde{Q} = JRJ^T = \begin{pmatrix} Q_1 & & & \\ & Q_2 & & 0 \\ & & \ddots & \\ & 0 & & Q_n \end{pmatrix} \tag{53}$$

where

$$Q_i = \begin{pmatrix} e_{ri}^2 \cos^2\phi_i + e_{\phi i}^2 r_i^2 \sin^2\phi_i & (e_{ri}^2 - e_{\phi i}^2 r_i^2)\sin\phi_i \cos\phi_i \\ (e_{ri}^2 - e_{\phi i}^2 r_i^2)\sin\phi_i \cos\phi_i & e_{ri}^2 \sin^2\phi_i + e_{\phi i}^2 r_i^2 \cos^2\phi_i \end{pmatrix} \tag{54}$$

For a numerical example, we take the ten points listed in Table 1 as observations and assume that their standard errors are

$$e_{ri} = 0.048, \; e_{\phi i} = 27.5°, \; i = 1, 2, \ldots, n. \tag{55}$$

We made three adjustments. First, the r,ϕ-data were used together with the model Eqs. (47). In the second adjustment, the x,y-data were used together with the model Eqs. (49) and the transformation function (48) in a utility program[15] based on the normal Eqs. (37). The results of both adjustments were identical, as they should be, and they are

TABLE 1. OBSERVATIONS ϕ AND r AND CORRESPONDING CARTESIAN COORDINATES

ϕ	r	x	y
206.6°	0.559	-0.50	-.025
26.6°	1.342	1.20	0.60
26.6°	2.236	2.00	1.00
26.6°	3.354	3.00	1.50
26.6°	4.472	4.00	2.00
123.7°	1.803	-1.00	1.50
92.9°	1.952	-0.10	1.95
68.2°	2.693	1.00	2.50
52.4°	4.100	2.50	3.25
42.0°	6.727	5.00	4.50

TABLE 2. ADJUSTMENT RESULTS

Case 1 and 2 (Original and Transformed Problem)

$a = 0.381 \pm 0.298$

$b = 1.141 \pm 0.744$ $\qquad c_{ab} = 0.015065$

$m_o = 1.24541$

Case 3 (Falsified Problem)

$a = 0.680 \pm 0.407$

$b = 1.837 \pm 0.259$ $\qquad c_{ab} = -0.568659$

$m_o = 1.75646$

The standard error of weight one, m_o, is not included in the standard errors of the parameters.

listed in Table 2. The listed standard errors of the parameters are the square roots of the diagonal elements of V_t, computed with formula (36). The correlation coefficient c_{ab} is the off-diagonal element of the correlation matrix C_t, defined by

$$C_t = D_t^{-1/2} V_t D_t^{-1/2} \tag{56}$$

where D_t is the diagonal matrix of V_t. The standard error of weight one is defined by

$$m_o = \left[\frac{1}{n-p} c^T R^{-1} c\right]^{-1/2} = \left[\frac{1}{n-p} W\right]^{1/2} . \tag{57}$$

Figure 1a shows the result of the adjustment in the ϕ,r-plane, i.e., in the plane of the original observations. The accuracies of the observations are indicated by error ellipses around the observed points. The adjustment is indicated by connecting the observed points with the corresponding corrected locations on the fitted curve. The figure shows that all adjustments are in the direction of largest uncertainties.

Figure 1b shows the same result in the x,y-plane. The accuracies of the transformed observations are again indicated by error ellipses, corresponding to the transformed variance-covariance matrices $\tilde{Q}_i$. In this presentation the adjustments seem to be in directions other than those with largest uncertainties. This is typical for nonlinear transformations of observations. The object of the fitting is to minimize residuals of the original observations. The presentation in the x,y-plane is distorted by the nonlinearity of the transformation.

In a third adjustment we used the x,y-data, the model Eq. (49), and the variance-covariance matrix $\tilde{Q}$, defined by Eq. (53). The treatment, suggested by Deming[1] and other authors, was described in Section III, Eqs. (44) through (46), as a falsification of the problem. The numerical results of this adjustment are listed in Table 2. They are different from the previous results, and the increase of m_o indicates that the solution is not optimal. We notice also that the correlation coefficient c_{ab} has changed its magnitude and sign.

Figure 2b shows the results of the adjustment in the x,y-plane. It indicates that the adjustment would indeed be optimal, if x,y were the observations and $\tilde{Q}$ were their variance-covariance matrix. However, when the same results are plotted in the ϕ,y-plane, Figure 2a, then it becomes obvious that the adjustment has not achieved the goal to minimize the residuals of the original observations ϕ,r. The treatment of

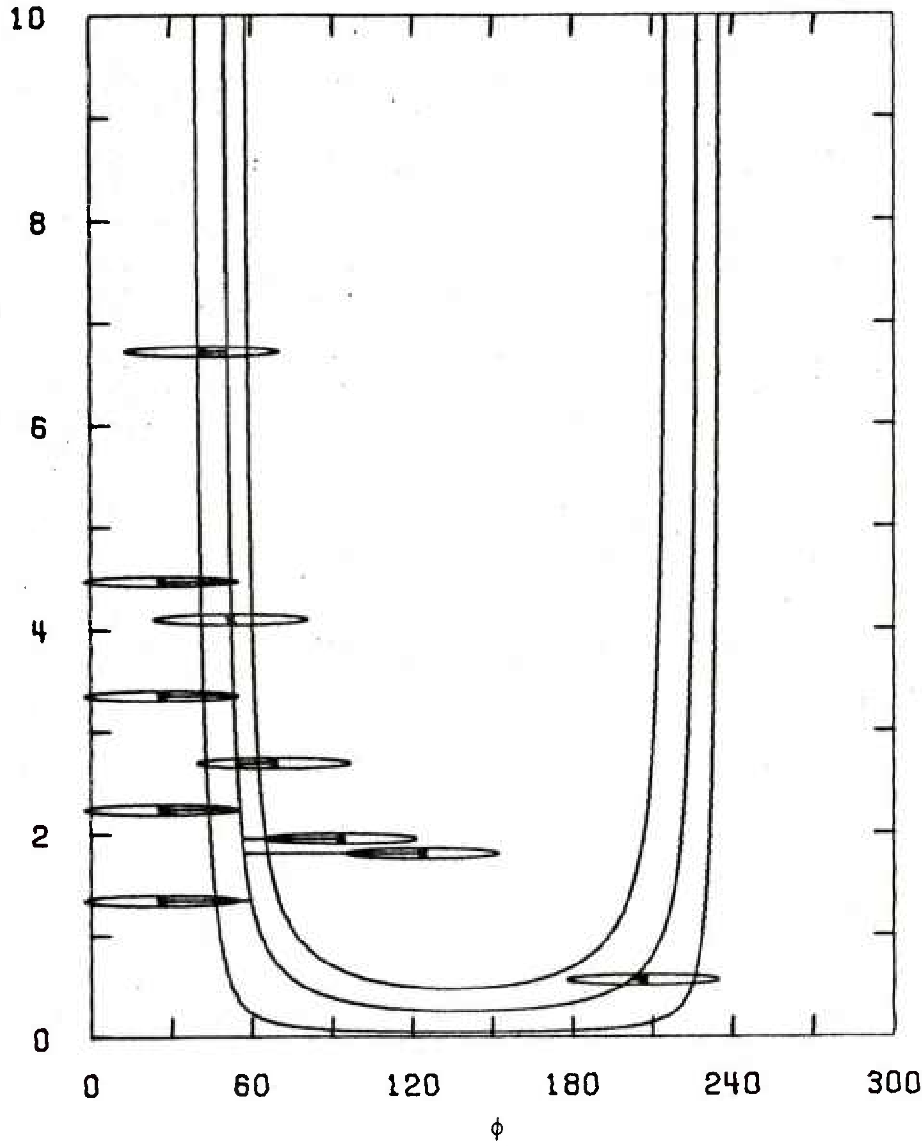

Figure 1a. Adjustment in ϕ,r-space.

The data are shown with their one standard error ellipses and the adjusted curve is shown with one standard error confidence limits. The same results are shown by Figure 1b in the cartesian x,y-plane.

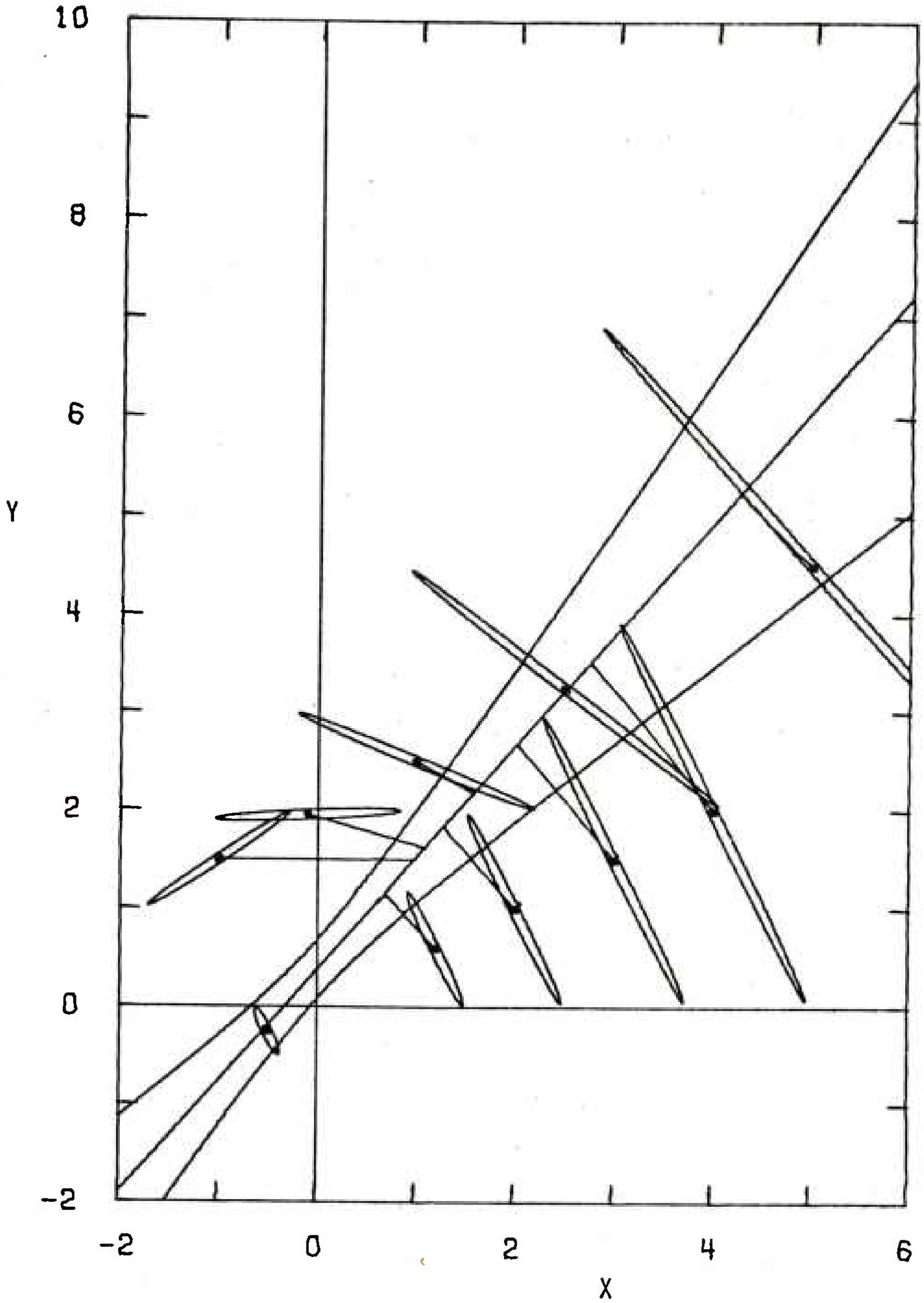

Figure 1b. Adjustment in ϕ,r-space.

The transformed data are shown with their one standard error ellipses and the adjusted line is shown with one standard error confidence limits. The same results are shown by Figure 1a in the ϕ,r-plane of observations.

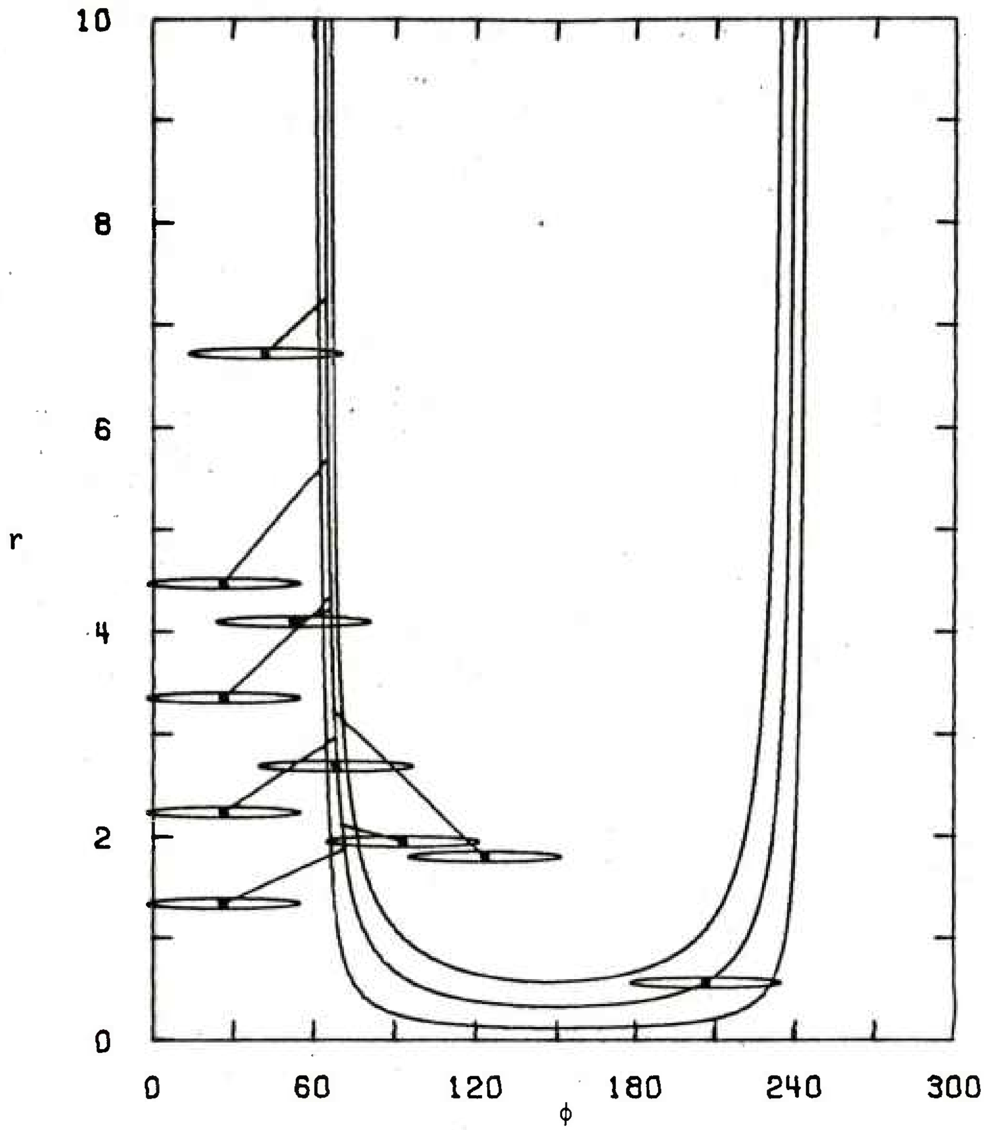

Figure 2a. Falsified Adjustment in x,y-space.

The data are shown with their one standard error ellipses and the adjusted curve is shown with one standard error confidence limits. The same results are shown in Figure 2b in the cartesian x,y-plane.

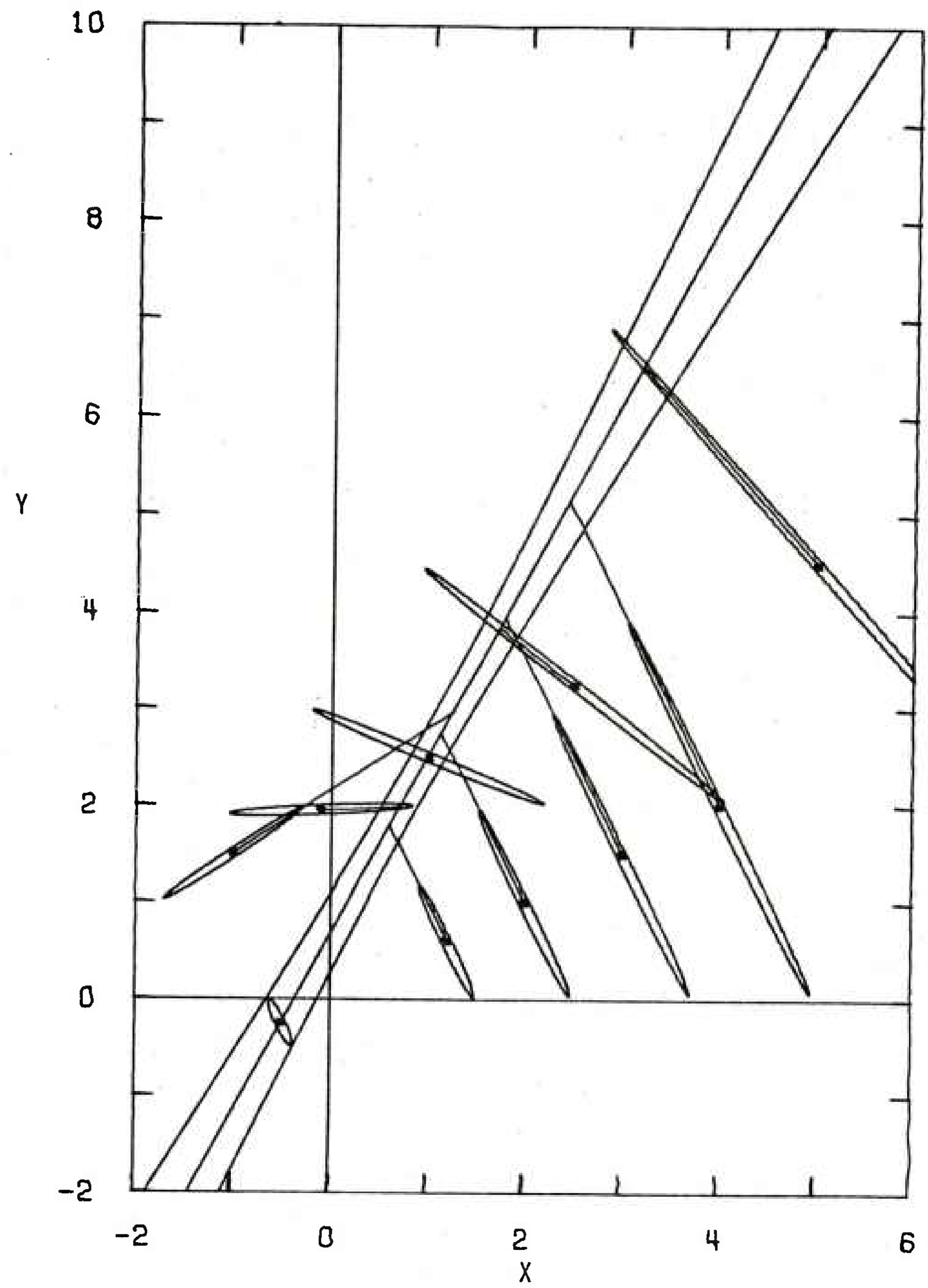

Figure 2b. Falsified Adjustment in x,y-space.

The transformed data are shown with their one standard error ellipses and the adjusted line is shown with one standard error confidence limits. The same results are shown in Figure 2a in the ϕ,r-plane of observations.

transformations of observations in this form is a falsification of the problem. The results are approximations to the least squares solution, but since the quality of the approximations is not known, they may be useful only as initial approximations for a least squares algorithm. However, in a case like this example, an initial approximation could be simpler obtained, e.g., graphically by drawing a straight line in the x,y-plane through the observations.

Next, we present an example for the linearization of parameters. Let the model equation be

$$y - Ax^B \exp\left(\frac{C}{x}\right) = 0 , \tag{58}$$

where x and y are observations and A, B, and C are model parameters. An equivalent model formulation is

$$\ln y - a - b \ln x - \frac{c}{x} = 0 . \tag{59}$$

In Eq. (59) the parameters a, b, and c enter linearly. One can expect a much better performance of solution algorithms if Eq. (59) is used. The parameter transformation is in this example

$$\begin{aligned} A &= e^a , \\ B &= b , \\ C &= c , \end{aligned} \tag{60}$$

and the Jacobian matrix, needed in Eq. (23) is

$$\frac{\partial(A,B,C)}{\partial(a,b,c)} = \begin{pmatrix} e^a & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} . \tag{61}$$

Another example is the trigonometric model

$$y - A \cos \frac{x-B}{C} = 0 . \tag{62}$$

An equivalent model is

$$y - a \sin(cx) - b \cos(cx) \; . \tag{63}$$

The corresponding parameter transformation is

$$\left.\begin{aligned} a &= A \sin(B/C) \; , \\ b &= A \cos(B/C) \; , \\ c &= 1/C \; , \end{aligned}\right\} \tag{64}$$

with the Jacobian matrix

$$\frac{\partial(A,B,C)}{\partial(a,b,c)} = \left(\frac{\partial(a,b,c)}{\partial(A,B,C)}\right)^{-1} =$$

$$= \begin{pmatrix} \sin(B/C) & (A/C)\cos(B/C) & -(AB/C^2)\cos(B/C) \\ \cos(B/C) & -(A/C)\sin(B/C) & (AB/C^2)\sin(B/C) \\ 0 & 0 & -1/C^2 \end{pmatrix}^{-1} \; , \tag{65}$$

In this example, the model (63) is linear only with respect to two parameters. However, the difference between numerical treatments of the problem is dramatic if one uses Eq. (62) or Eq. (63), respectively. In numerical experiments we found that in order to achieve convergence, one had to start with parameter values A,B,C that were very close to their least squares values. Using the parameters a,b,c and the model Eq. (63), one achieves fast convergence, e.g., with the initial values a=b=0.

V. SUMMARY AND CONCLUSIONS

Manipulations of model equations that produce simpler but equivalent equations can greatly facilitate the preparation of least squares problems (e.g., computer programming) for utility routines. The manipulations can also improve the performance of numerical algorithms. If the manipulations are merely algebraic and/or involve nonlinear transformations of the model parameters, then their application is straightforward and their implementation simple. If, however, the manipulations include transformations of observations, then one has to transform also the normal equations correspondingly. Neglect of this transformation falsifies the problem and produces results that are of unknown quality and equally reliable as, e.g., a graphical construction of a fitting curve. A correct implementation of transformations of observations requires the programming of the transformation function, including its first and second order derivatives. It also does not improve the performances of algorithms. Therefore, in most cases, it is more efficient to formulate the model equations in terms of the original observations, thereby avoiding the programming of the transformation function.

The need for second order derivatives of the model equations has been often overlooked. In order to avoid the programming of these derivatives, most authors suggest to use a first order Gauss-Newton algorithm for the solution of the normal equations, instead of a second-order Newton-Raphson algorithm. The performance of the former may be often comparable to the latter, because even with more iterations, the computing effort can be less due to the simpler equations. Second-order derivatives of the model equations (and of the transformation function) are, however, needed to compute the linear terms in formulas for variance estimates of the results. Their neglect cannot be justified cursorily by the argument that linearized model equations are already second order accurate and, therefore, their second order derivatives are not needed. It can be shown that the linearized normal equations do contain these derivatives and, therefore, are needed in the linearized variance propagation formula. Formulas for variance estimates that do not contain second order derivatives are less than first order accurate.

REFERENCES

1. W. Edwards Deming, Statistical Adjustment of Data, John Wiley and Sons, London, 1944.

2. John E. Freund, Mathematical Statistics, Prentice Hall, Englewood Cliffs, NJ, 1962.

3. Karel Rektorys, ed., Survey of Applicable Mathematics, The M.I.T. Press, Cambridge, MA, 1969.

4. Yonathan Bard, Nonlinear Parameter Estimation, Academic Press, New York, NY, 1974.

5. Yo Lun Chou, Statistical Analysis, Holt, Rinehart, and Winston, New York, NY, 1975.

6. A.S.C. Ehrenberg, Data Reduction, John Wiley and Sons, New York, NY, 1975.

7. Thomas H. Wonnacott and Ronald F. Wonnacott, Introductory Statistics, 3rd ed., John Wiley and Sons, New York, NY, 1977.

8. R.M. Passi, "Use of Nonlinear Least Squares in Meteorological Applications", Journal of Applied Meteorology, Vol. 16, pp. 827-832, 1977; and Vol. 17, pp. 1579-1580, 1978.

9. Heike von Benda, "Zur Gitterkonstantenbestimung mit Ausgleichsmetoden", Zeitschrift für Kristallographie, Vol. 149, pp. 205-209, 1979.

10. Ph. Wahl, "Analysis of Fluorescence Anisotrophy Decays by a Least Square Method", Biophysical Chemistry, Vol. 10, pp. 91-104, 1973.

11. H.J. Britt and R.H. Luecke, "The Estimation of Parameters in Nonlinear, Implicit Models", Technometrics, Vol. 15, pp. 233-247, 1973.

12. P.A.D. DeMaine, "Automatic Curve Fitting, I. Test Methods", Computers and Chemistry, Vol. 2, pp. 1-6, 1978.

13. Allen J. Pope, "Two Approaches to Nonlinear Least Squares Adjustments", The Canadian Surveyor, Vol. 28, pp. 663-669, 1974.

14. Robert E. Barieau and B.J. Dalton, "Nonlinear Regression and the Principle of Least Squares", Bureau of Mines Report of Investigations 6900, 1967.

15. Aivars Celmiņš, "A Manual for General Least Squares Model Fitting", Ballistic Research Laboratory Technical Report ARBRL-TR-02167, 1979. (AD #B040229L)

16. E. Stark and E. Mikhail, "Least Squares and Non-Linear Functions", Photogrammetric Engineering, pp. 405-412, 1973.

17. William H. Sachs, "Implicit Multifunctional Nonlinear Regression Analysis", Technometrics, Vol. 18, pp. 161-173, 1976.

18. Aivars Celmiņš, "Least Squares Adjustment with Finite Residuals for Non-Linear Constraints and Partially Correlated Data", Ballistic Research Laboratory Report BRL-R-1658, 1973. (AD #766283)

APPENDIX

DERIVATION OF ITERATION FORMULAS

We provide a set of iteration formulas that are derived from the Newton equation (34) by algebraic manipulations. First, we define the following matrices:

$$G = (F_X R F_X^T)^{-1} \tag{A-1}$$

$$A = R F_X^T G F_X - I \tag{A-2}$$

$$\Gamma = \left[I + AR(K^T F)_{XX}\right]^{-1} \tag{A-3}$$

$$E_o = \Gamma \cdot \left[AC - RF_X^T G F_X\right] \tag{A-4}$$

$$E_1 = \Gamma \cdot \left[RF_X^T G F_t + AR(K^T F)_{Xt}\right] \tag{A-5}$$

$$D_o = (K^T F)_{tX} - F_t^T G F_X R (K^T F)_{XX} \tag{A-6}$$

$$D_1 = (K^T F)_{tt} - F_t^T G F_X R (K^T F)_{Xt} \tag{A-7}$$

$$N = F_t^T G F_t - D_1 + D_0 E_1 \tag{A-8}$$

The iteration equations are

$$N\tau = F_t^T G (F_X C - F) + D_0 E_0 \tag{A-9}$$

$$K + \kappa = G(F_X C - F) + G\left[F_t + F_X R (K^T F)_{Xt}\right]\tau - G F_X R (K^T F)_{XX} \varepsilon \tag{A-10}$$

$$\varepsilon = E_0 - E_1 \tau \quad . \tag{A-11}$$

Numerical experiments have shown that the convergence of the iteration is enhanced if the equations are used in a subiteration mode by iterating alternatively on the parameters and residuals, respectively. For parameter subiteration only equations (A-9) and (A-10) are used, assuming $\varepsilon \equiv 0$. For residual subiteration one sets $\tau \equiv 0$ and uses equations (A-10) and (A-11).

In the variance formula (36) one uses N, defined by equation (A-8) and

$$S = F_t^T G F_X + D_0 \Gamma A. \tag{A-12}$$

Another equivalent set of Newton-Raphson iteration equations is given in Reference 13. None of the sets is numerically superior to the other, and both require subiterations of parameters and residuals for efficiency.

Gauss-Newton iteration equations can be obtained from Newton-Raphson iteration equations by setting all second order derivatives equal to zero. The convergence of Gauss-Newton algorithms is inferior, but in some applications they have a larger domain of convergence.

Iteration equations for least squares problems with transformations of observations can be obtained from the formulas in this Appendix by substituting

Q for R

Δ for C

and

Ξ for $(K^T F)_{XX}$.

Expressions for Q, Δ, and Ξ in terms of the model and the transformation functions are given in Section III, equations (41), (42), and (43).

DISTRIBUTION LIST

No. of Copies	Organization
12	Commander Defense Documentation Center ATTN: DDC-DDA Cameron Station Alexandria, VA 22314
1	Director Defense Advanced Research Projects Agency ATTN: Technical Library 1400 Wilson Boulevard Arlington, VA 22209
1	Director Defense Civil Preparedness Agency ATTN: Technical Library Washington, DC 20301
7	Director Defense Nuclear Agency ATTN: RATN SPSS STSI/Archives STSP STTL/Tech Lib (2 cys) STVL Washington, DC 20305
1	Director US Army BMD Advanced Technology Center P.O. Box 1500, W. Station Huntsville, AL 35807
1	Commander US Army Ballistic Missile Defense Systems Command Huntsville, AL 35804
1	Commander US Army Engineer Waterways Experiment Station ATTN: Technical Library P.O. Box 631 Vicksburg, MS 39180
1	Commander US Army Materiel Development and Readiness Command ATTN: DRCDMD-ST 5001 Eisenhower Avenue Alexandria, VA 22333
1	Commander US Army Materiel Development and Readiness Command ATTN: Analysis & Computational Division/Mr. Sanchez White Sands Missile Range, NM 88002
2	Commander US Army Armament Research and Development Command ATTN: DRDAT-TSS (2 Cys) Dover, NJ 07801
2	Commander US Army Armament Research and Development Command ATTN: DRDAR-LCN/P. Angelotti Dover, NJ 07801
1	Commander US Army Armament Materiel Readiness Command ATTN: DRSAR-LEP-L/Tech Lib Rock Island, IL 61299
1	Director USA ARRADCOM Benet Weapons Laboratory ATTN: DRDAR-LCB-TL Watervliet, NY 12189
1	Commander US Army Aviation Research and Development Command ATTN: DRDAV-E 4300 Goodfellow Blvd. St. Louis, MO 63120

DISTRIBUTION LIST

No. of Copies	Organization
1	Director US Army Air Mobility Research and Development Laboratory Ames Research Center Moffett Field, CA 94035
1	Commander US Army Communications Research and Development Command ATTN: DRDCO-PPA-SA Fort Monmouth, NJ 07703
1	Commander US Army Electronics Research and Development Command Technical Support Activity ATTN: DELSD-L Fort Monmouth, NJ 07703
6	Commander US Army Harry Diamond Labs ATTN: DRXDO-TI DRXDO-TI/012 DRXDO-NP DRXDO-RBH/Mr. Caldwell DELHD-RBA/J. Rosado DELMD-ASE 2800 Powder Mill Road Adelphi, MD 20783
1	Commander US Army Missile Research and Development Command ATTN: DRDMI-R Redstone Arsenal, AL 35809
1	Commander US Army Missile Research and Development Command ATTN: DRDMI-YDL Redstone Arsenal, AL 35809
1	Commander US Army Missile Research and Development Command ATTN: DRDMI-RDK Redstone Arsenal, AL 35809
2	Commander US Army Tank Automotive Research and Development Command ATTN: DRDTA-UL DRDTA-ZSS/T. Szako Warren, MI 48090
2	Commander US Army Nuclear Agency ATTN: ACTA-NAW Technical Library 7500 Backlick Road, Bldg. 2073 Springfield, VA 22150
1	Director US Army TRADOC Systems Analysis Activity ATTN: ATAA-SL/Tech Lib White Sands Missile Range, NM 88002
1	Commander US Army Jefferson Proving Ground ATTN: STEJP-TD-D Madison, IN 47250
1	Commander US Army Materials and Mechanics Research Center ATTN: DRXMR-ATL Watertown, MA 02172
1	Commander US Army Concepts Analysis Agency 8120 Woodmont Avenue Bethesda, MD 20014

DISTRIBUTION LIST

No. of Copies	Organization
1	Director Strategic Systems Project Ofc ATTN: NSP-43/Tech Lib Department of the Navy Washington, DC 20360
3	Commander US Naval Air Systems Command ATTN: AIR-604 Washington, DC 20360
3	Commander US Naval Ordnance Systems Command ATTN: ORD-9132 Washington, DC 20360
2	Commander and Director David W. Taylor Naval Ship Research and Development Ctr ATTN: Tech Lib Aerodynamic Lab Bethesda, MD 20084
1	Commander US Naval Ordnance Station ATTN: Code FS13A/P. Sewell Indian Head, MD 20640
1	Commander US Naval Research Laboratory ATTN: Tech Info Div Washington, DC 20375
2	Commander US Naval Surface Weapons Ctr ATTN: Code WX21/Tech Lib Code 240/C.J. Aronson Silver Spring, MD 20910
1	Commander US Naval Weapons Center ATTN: Code 553/Tech Lib China Lake, CA 93555

No. of Copies	Organization
1	AFRPL/LKCB(Dr. Horning) Edwards AFB, CA 93523
1	AFATL(Tech Lib) Eglin AFB, FL 32542
1	AFWL/DE-I Kirtland AFB, NM 87117
1	AFWL/DEX Kirtland AFB, NM 87117
1	AFWL/SUL Kirtland AFB, NM 87117
1	AFWL/R.Henny Kirtland AFB, NM 87117
1	ASD/XRA(Stinfo) Wright-Patterson AFB, OH 45433
1	National Oceanic and Atmospheric Admin (C121) ATTN: A.F. Pope Rockville, MD 20852
1	National Center for Atmospheric Research ATTN: R.M. Passi Boulder, CO 80803
3	National Bureau of Standards ATTN: D. Shier D. Kahaner W.S. Horton (561) Washington, DC 20234
1	US Geological Survey ATTN: J.R. Fisher Washington, DC 20242
1	National Heart and Lung Institute Biometrics Research Branch ATTN: J.H. Ware Bethesda, MD 20014

DISTRIBUTION LIST

No. of Copies	Organization
1	National Institute of Health ATTN: R.I. Shrafer, Bldg. 12A, Rm2055 900 Rockville Pike Bethesda, MD 20205
1	Director Brookhaven National Laboratory Associate Universities, Inc. ATTN: W.H. Sachs Building 801 Upton, Long Island, NY 11973
1	Director Lawrence Livermore National Laboratory ATTN: Tech Lib P.O. Box 969 Livermore, CA 94550
1	Director Lawrence Livermore National Laboratory ATTN: Tech Info Dept L-3 P.O. Box 808 Livermore, CA 94550
1	Director Los Alamos Scientific Lab ATTN: Doc Contr for Rpts Lib P.O. Box 1663 Los Alamos, NM 87544
1	Director Oak Ridge National Laboratory Nuclear Division ATTN: D.J.W. Longworth P.O. Box Y Oak Ridge, TN 37830
1	Sandia Laboratories ATTN: Doc Contr for 3141 Sandia Rpt Collection P.O. Box 5800 Albuquerque, NM 87115
1	Director Jet Propulsion Laboratory ATTN: Tech Lib 2800 Oak Grove Drive Pasadena, CA 91103
1	Director National Aeronautics and Space Administration Langley Research Center ATTN: MS 185/Tech Lib Langely Station Hampton, VA 23365
2	Director National Aeronautics and Space Administration George C. Marshall Space Flight Center ATTN: MS-I/Lib R-AERO-AE/A. Felix Huntsville, AL 35812
1	Director NASA Scientific and Technical Information Facility ATTN: SAK/DL P.O. Box 8757 BWI Airport, MD 21240
1	Aerospace Corporation Performance Analysis Department ATTN: J.T. Betts El Segundo, CA 90245
1	ARO, Inc. ATTN: Tech Lib Arnold AFS, TN 37389
1	The Boeing Company ATTN: Aerospace Library P.O. Box 2707 Seattle, WA 98124

DISTRIBUTION LIST

No. of Copies	Organization
2	E.I. Du Pont de Nemours Co. Engineering Department ATTN: R.D. Snee C.H. White Wilmington, DE 19898
1	General Electric Company-TEMPO ATTN: E. Bryant 220 S. Main Street, Rm 206 Bel Air, MD 21014
1	General Electric Company Armament Department ATTN: W. Hathaway Lakeside Avenue Burlington, VT 05401
1	General Electric Company-TEMPO ATTN: DASIAC P.O. Drawer QQ Santa Barbara, CA 93102
1	Pacific Sierra Research Corp. ATTN: H. Brode 1456 Cloverfield Boulevard Santa Monica, CA 90404
1	Physics International Corp. ATTN: Tech Lib 2700 Meced Street San Leandro, CA 94577
1	R&D Associates ATTN: Tech Lib P.O. Box 9695 Marina del Rey, CA 90291
1	RCA Corporation David Sarnoff Research Center ATTN: Head, Applied Math & Physical Sciences Princeton, NJ 08540

No. of Copies	Organization
1	Science Applications, Inc. ATTN: Tech Lib P.O. Box 2351 LaJolla, CA 92037
1	Systems, Science, & Software ATTN: Tech Lib P.O. 1620 LaJolla, CA 90237
1	Terra Tek, Inc. ATTN: Tech Lib 420 Wakara Way Salt Lake City, UT 84108
1	TRW Systems Group ATTN: Tech Info Ctr/S-1930 One Space Park Redonodo Beach, CA 92078
1	Union Carbide Corporation Chemicals and Plastics ATTN: H.J. Britt P.O. Box 8361 Charleston, WV 25303
1	California Institute of Tech Guggenheim Aeronautical Lab ATTN: Tech Lib Pasadena, CA 91104
1	California State College ATTN: R.F. Dennemeyer San Bernardino, CA 92407
1	Colorado State University Department of Mathematics ATTN: G.D. Taylor Fort Collins, CO 80523

DISTRIBUTION LIST

No. of Copies	Organization
1	University of Colorado Computer Science Department ATTN: R. Schnabel Boulder, CO 80304
1	Cornell University Computer Science Department ATTN: J. Dennis Ithaca, NY 14850
1	University of Delaware College of Arts and Sciences ATTN: M.Z. Nashed 501 Kirkbride Ofc Building Newark, DE 19711
1	Franklin Institute ATTN: Dr. Carfagno Race & 20th Streets Philadelphia, PA 19103
1	The George Washington Univ. Department of Operations Res. ATTN: A.V. Fiacco Washington, DC 20234
1	The Johns Hopkins University Applied Physics Laboratory Johns Hopkins Road Laurel, MD 20810
1	The Johns Hopkins University Mathematical Sciences ATTN: R.H. Byrd Baltimore, MD 21218
1	University of Maryland Baltimore Campus Mathematics Department ATTN: P. Kumar Baltimore, MD 21228

No. of Copies	Organization
1	University of Maryland ATTN: G.W. Stewart College Park, MD 20742
1	University of Maryland Department of Mathematics ATTN: G. Young College Park, MD 20742
1	Massachusetts Institute of Technology Department of Aeronautics and Astronautics ATTN: Tech Lib 77 Massachusetts Avenue Cambridge, MA 02139
1	Ohio State University Department of Aeronautics and Astronautical Engineering ATTN: Tech Lib Columbus, OH 43210
1	Ohio State University Department of Geodetic Science ATTN: R.H. Rapp 1958 Neil Avenue Columbus, OH 43210
1	Pennsylvania State University Computer Sciences Department ATTN: P.A.D. DeMaine University Park, PA 16807
1	Polytechnic Institute of Brooklyn Graduate Center ATTN: Tech Lib Farmingdale, NY 11735
2	Southwest Research Institute ATTN: W.E. Baker A.B. Wenzel 8500 Culebra Road San Antonio, TX 78206

DISTRIBUTION LIST

No. of Copies	Organization
1	Stanford University ATTN: G.H. Golub Stanford, CA 94305
1	Texas A&M University ATTN: S.B. Childs College Station, TX 77843
1	The University of Texas ATTN: D.R. Barker 5323 Harry Hives Boulevard Dallas, TX 75234
4	University of Wisconsin Mathematics Research Center ATTN: N. Draper P. Newbold D. Gay C. deBoor 610 Walnut Street Madison, WI 53706
	Aberdeen Proving Ground Dir, USAMSAA ATTN: DRXSY-D DRXSY-MP, H. Cohen Cdr, USATECOM ATTN: DRSTE-TO-F Dir, USACSL, Bldg. E3516, EA ATTN: DRDAR-CLB-PA